



OIPE

ENTERED

RAW SEQUENCE LISTING

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:55

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

```

3 <110> APPLICANT: ASAKO, HIROYUKI
4     MATSUMURA, KENJI
5     SHIMIZU, MASATOSHI
6     ITO, NOBUYA
7     WAKITA, RYUHEI
9 <120> TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
10    4-HALO-3-HYDROXYBUTANOATE
12 <130> FILE REFERENCE: 7372-72249
14 <140> CURRENT APPLICATION NUMBER: 10/004,115A
C--> 15 <141> CURRENT FILING DATE: 2002-10-23
17 <150> PRIOR APPLICATION NUMBER: JP 2000-372704
18 <151> PRIOR FILING DATE: 2000-12-07
20 <150> PRIOR APPLICATION NUMBER: JP 2001-006144
21 <151> PRIOR FILING DATE: 2001-01-15
23 <150> PRIOR APPLICATION NUMBER: JP 2001-026594
24 <151> PRIOR FILING DATE: 2001-02-02
26 <150> PRIOR APPLICATION NUMBER: JP 2001-175175
27 <151> PRIOR FILING DATE: 2001-06-11
29 <160> NUMBER OF SEQ ID NOS: 37
31 <170> SOFTWARE: PatentIn Ver. 2.1
33 <210> SEQ ID NO: 1
34 <211> LENGTH: 325
35 <212> TYPE: PRT
36 <213> ORGANISM: Penicillium citrinum
38 <400> SEQUENCE: 1
39 Met Ser Asn Gly Lys Thr Phe Thr Leu Ser Asn Gly Val Lys Ile Pro
40   1           5           10           15
42 Gly Val Gly Phe Gly Thr Phe Ala Ser Glu Gly Ser Lys Gly Glu Thr
43   20           25           30
45 Tyr Thr Ala Val Thr Thr Ala Leu Lys Thr Gly Tyr Arg His Leu Asp
46   35           40           45
48 Cys Ala Trp Tyr Tyr Leu Asn Glu Gly Glu Val Gly Glu Gly Ile Arg
49   50           55           60
51 Asp Phe Leu Lys Glu Asn Pro Ser Val Lys Arg Glu Asp Ile Phe Val
52   65           70           75           80
54 Cys Thr Lys Val Trp Asn His Leu His Arg Tyr Glu Asp Val Leu Trp
55   85           90           95
57 Ser Ile Asp Asp Ser Leu Lys Arg Leu Gly Leu Asp Tyr Val Asp Met
58   100          105          110
60 Phe Leu Val His Trp Pro Ile Ala Glu Lys Asn Gly Gln Gly Glu
61   115          120          125
63 Pro Lys Ile Gly Pro Asp Gly Lys Tyr Val Ile Leu Lys Asp Leu Thr
64   130          135          140

```

RAW SEQUENCE LISTING

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:55

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

```

66 Glu Asn Pro Glu Pro Thr Trp Arg Ala Met Glu Lys Ile Tyr Glu Asp
67 145                               150                               155                               160
69 Arg Lys Ala Arg Ser Ile Gly Val Ser Asn Trp Thr Ile Ala Asp Leu
70                               165                               170                               175
72 Glu Lys Met Ser Lys Phe Ala Lys Val Met Pro His Ala Asn Gln Ile
73                               180                               185                               190
75 Glu Ile His Pro Phe Leu Pro Asn Glu Glu Leu Val Gln Tyr Cys Phe
76                               195                               200                               205
78 Ser Lys Asn Ile Met Pro Val Ala Tyr Ser Pro Leu Gly Ser Gln Asn
79 210                               215                               220
81 Gln Val Pro Thr Thr Gly Glu Arg Val Ser Glu Asn Lys Thr Leu Asn
82 225                               230                               235                               240
84 Glu Ile Ala Glu Lys Gly Gly Asn Thr Leu Ala Gln Val Leu Ile Ala
85                               245                               250                               255
87 Trp Gly Leu Arg Arg Gly Tyr Val Val Leu Pro Lys Ser Ser Asn Pro
88                               260                               265                               270
90 Lys Arg Ile Glu Ser Asn Phe Lys Ser Ile Glu Leu Ser Asp Ala Asp
91 275                               280                               285
93 Phe Glu Ala Ile Asn Ala Val Ala Lys Gly Arg His Phe Arg Phe Val
94 290                               295                               300
96 Asn Met Lys Asp Thr Phe Gly Tyr Asp Val Trp Pro Glu Glu Thr Ala
97 305                               310                               315                               320
99 Lys Asn Leu Ser Ala
100                               325
103 <210> SEQ ID NO: 2
104 <211> LENGTH: 978
105 <212> TYPE: DNA
106 <213> ORGANISM: Penicillium citrinum
108 <220> FEATURE:
109 <221> NAME/KEY: CDS
110 <222> LOCATION: (1)..(975)
112 <400> SEQUENCE: 2
113 atg tct aac gga aag act ttc aca ttg agc aac ggc gtc aag att cct 48
114 Met Ser Asn Gly Lys Thr Phe Thr Leu Ser Asn Gly Val Lys Ile Pro
115 1 5 10 15
117 ggc gtc ggc ttt ggt acc ttc gct agt gaa ggt tcc aag ggc gag acc 96
118 Gly Val Gly Phe Gly Thr Phe Ala Ser Glu Gly Ser Lys Gly Glu Thr
119 20 25 30
121 tat act gct gtc acc act gcc ctg aag acc ggt tac cgt cac ttg gac 144
122 Tyr Thr Ala Val Thr Thr Ala Leu Lys Thr Gly Tyr Arg His Leu Asp
123 35 40 45
125 tgt gcc tgg tac tac ctg aac gag ggt gag gtt ggt gag ggt atc cgt 192
126 Cys Ala Trp Tyr Tyr Leu Asn Glu Gly Glu Val Gly Glu Gly Ile Arg
127 50 55 60
129 gac ttc ctg aag gag aac ccc tcg gtg aag cgt gag gac atc ttc gtc 240
130 Asp Phe Leu Lys Glu Asn Pro Ser Val Lys Arg Glu Asp Ile Phe Val
131 65 70 75 80
133 tgc acc aag gtg tgg aac cac ctc cac cgt tat gag gac gtc ctc tgg 288
134 Cys Thr Lys Val Trp Asn His Leu His Arg Tyr Glu Asp Val Leu Trp

```

RAW SEQUENCE LISTING

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:55

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

| | | | | | | | |
|-----|---|-----|-----|-----|-----|-----|--|
| 135 | | 85 | | 90 | | 95 | |
| 137 | tcc att gac gac tcc ctg aag cgt ctt gga ctt gac tac gtt gat atg | 336 | | | | | |
| 138 | Ser Ile Asp Asp Ser Leu Lys Arg Leu Gly Leu Asp Tyr Val Asp Met | | | | | | |
| 139 | | 100 | | 105 | | 110 | |
| 141 | ttc ctc gtt cac tgg ccc att gct gcc gag aag aat ggc cag ggt gag | 384 | | | | | |
| 142 | Phe Leu Val His Trp Pro Ile Ala Ala Glu Lys Asn Gly Gln Gly Glu | | | | | | |
| 143 | | 115 | | 120 | | 125 | |
| 145 | ccc aag att ggc cct gac ggc aaa tac gtc att ctc aag gac ctg acc | 432 | | | | | |
| 146 | Pro Lys Ile Gly Pro Asp Gly Lys Tyr Val Ile Leu Lys Asp Leu Thr | | | | | | |
| 147 | | 130 | | 135 | | 140 | |
| 149 | gag aac ccc gag ccc aca tgg cgc gct atg gag aag att tat gag gat | 480 | | | | | |
| 150 | Glu Asn Pro Glu Pro Thr Trp Arg Ala Met Glu Lys Ile Tyr Glu Asp | | | | | | |
| 151 | 145 | | 150 | | 155 | 160 | |
| 153 | cgc aag gcc agg tcc att ggt gtc tcc aac tgg acc att gcc gac ctt | 528 | | | | | |
| 154 | Arg Lys Ala Arg Ser Ile Gly Val Ser Asn Trp Thr Ile Ala Asp Leu | | | | | | |
| 155 | | 165 | | 170 | | 175 | |
| 157 | gag aag atg tcc aag ttc gcc aag gtc atg cct cac gcc aac cag atc | 576 | | | | | |
| 158 | Glu Lys Met Ser Lys Phe Ala Lys Val Met Pro His Ala Asn Gln Ile | | | | | | |
| 159 | | 180 | | 185 | | 190 | |
| 161 | gag att cac ccc ttc ctg ccc aac gag gag ctg gtg cag tac tgc ttc | 624 | | | | | |
| 162 | Glu Ile His Pro Phe Leu Pro Asn Glu Glu Leu Val Gln Tyr Cys Phe | | | | | | |
| 163 | | 195 | | 200 | | 205 | |
| 165 | tcc aag aac att atg ccc gtg gcc tac tct cct ctg ggc tcg cag aac | 672 | | | | | |
| 166 | Ser Lys Asn Ile Met Pro Val Ala Tyr Ser Pro Leu Gly Ser Gln Asn | | | | | | |
| 167 | | 210 | | 215 | | 220 | |
| 169 | cag gtt ccc acc acc ggt gag cgg gtc agc gag aac aag act ctg aac | 720 | | | | | |
| 170 | Gln Val Pro Thr Thr Gly Glu Arg Val Ser Glu Asn Lys Thr Leu Asn | | | | | | |
| 171 | 225 | | 230 | | 235 | 240 | |
| 173 | gag atc gcc gag aag ggc ggc aac acc ctt gct cag gtt ctt att gcc | 768 | | | | | |
| 174 | Glu Ile Ala Glu Lys Gly Gly Asn Thr Leu Ala Gln Val Leu Ile Ala | | | | | | |
| 175 | | 245 | | 250 | | 255 | |
| 177 | tgg ggt ctg cgc cgt ggc tac gtc gtt ctc ccc aag agc tcc aac ccc | 816 | | | | | |
| 178 | Trp Gly Leu Arg Arg Gly Tyr Val Val Leu Pro Lys Ser Ser Asn Pro | | | | | | |
| 179 | | 260 | | 265 | | 270 | |
| 181 | aag cgc att gag tcc aac ttc aag agc att gag ctc tcc gat gcc gac | 864 | | | | | |
| 182 | Lys Arg Ile Glu Ser Asn Phe Lys Ser Ile Glu Leu Ser Asp Ala Asp | | | | | | |
| 183 | | 275 | | 280 | | 285 | |
| 185 | ttt gaa gcc atc aat gcc gtt gcc aag ggt cgt cac ttc cgt ttc gtc | 912 | | | | | |
| 186 | Phe Glu Ala Ile Asn Ala Val Ala Lys Gly Arg His Phe Arg Phe Val | | | | | | |
| 187 | | 290 | | 295 | | 300 | |
| 189 | aac atg aag gat act ttc gga tat gat gtc tgg ccc gag gag acc gcc | 960 | | | | | |
| 190 | Asn Met Lys Asp Thr Phe Gly Tyr Asp Val Trp Pro Glu Glu Thr Ala | | | | | | |
| 191 | 305 | | 310 | | 315 | 320 | |
| 193 | aag aac ctg tct gcg tga | 978 | | | | | |
| 194 | Lys Asn Leu Ser Ala | | | | | | |
| 195 | | 325 | | | | | |
| 198 | <210> SEQ ID NO: 3 | | | | | | |
| 199 | <211> LENGTH: 17 | | | | | | |
| 200 | <212> TYPE: PRT | | | | | | |

RAW SEQUENCE LISTING

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:55

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

```

201 <213> ORGANISM: Penicillium citrinum
203 <400> SEQUENCE: 3
204 Asn Ile Met Pro Val Ala Tyr Ser Pro Leu Gly Ser Gln Asn Gln Val
205   1                               5               10               15
207 Pro
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 10
212 <212> TYPE: PRT
213 <213> ORGANISM: Penicillium citrinum
215 <400> SEQUENCE: 4
216 Ile Pro Gly Val Phe Gly Thr Phe Ala Ser
217   1                               5               10
220 <210> SEQ ID NO: 5
222 <400> SEQUENCE: 5
W--> 223 000
226 <210> SEQ ID NO: 6
227 <211> LENGTH: 14
228 <212> TYPE: PRT
229 <213> ORGANISM: Penicillium citrinum
231 <220> FEATURE:
232 <221> NAME/KEY: MOD_RES
233 <222> LOCATION: (6)..(7)
234 <223> OTHER INFORMATION: Variable amino acid
236 <400> SEQUENCE: 6
W--> 237 Tyr Glu Asp Val Leu Xaa Xaa Ile Asp Asp Ser Leu Lys Arg
238   1                               5               10
241 <210> SEQ ID NO: 7
243 <400> SEQUENCE: 7
W--> 244 000
247 <210> SEQ ID NO: 8
248 <211> LENGTH: 20
249 <212> TYPE: DNA
250 <213> ORGANISM: Artificial Sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
255 <400> SEQUENCE: 8
256 ggaacytgrt tytggswacc                                     20
259 <210> SEQ ID NO: 9
260 <211> LENGTH: 20
261 <212> TYPE: DNA
262 <213> ORGANISM: Artificial Sequence
264 <220> FEATURE:
265 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
267 <220> FEATURE:
268 <221> NAME/KEY: modified_base
269 <222> LOCATION: (3)
270 <223> OTHER INFORMATION: a, c, t, g, other or unknown
272 <220> FEATURE:
273 <221> NAME/KEY: modified_base

```

RAW SEQUENCE LISTING

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:55

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

274 <222> LOCATION: (6)
275 <223> OTHER INFORMATION: a, c, t, g, other or unknown
277 <220> FEATURE:
278 <221> NAME/KEY: modified_base
279 <222> LOCATION: (9)
280 <223> OTHER INFORMATION: a, c, t, g, other or unknown
282 <400> SEQUENCE: 9

W--> 283 tangcnacng gcataatatt 20
286 <210> SEQ ID NO: 10
287 <211> LENGTH: 20
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
294 <220> FEATURE:
295 <221> NAME/KEY: modified_base
296 <222> LOCATION: (3)
297 <223> OTHER INFORMATION: a, c, t, g, other or unknown
299 <220> FEATURE:
300 <221> NAME/KEY: modified_base
301 <222> LOCATION: (6)
302 <223> OTHER INFORMATION: a, c, t, g, other or unknown
304 <220> FEATURE:
305 <221> NAME/KEY: modified_base
306 <222> LOCATION: (9)
307 <223> OTHER INFORMATION: a, c, t, g, other or unknown
309 <400> SEQUENCE: 10

W--> 310 tangcnacng gcataatggt 20
313 <210> SEQ ID NO: 11
314 <211> LENGTH: 20
315 <212> TYPE: DNA
316 <213> ORGANISM: Artificial Sequence
318 <220> FEATURE:
319 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
321 <220> FEATURE:
322 <221> NAME/KEY: modified_base
323 <222> LOCATION: (3)
324 <223> OTHER INFORMATION: a, c, t, g, other or unknown
326 <220> FEATURE:
327 <221> NAME/KEY: modified_base
328 <222> LOCATION: (6)
329 <223> OTHER INFORMATION: a, c, t, g, other or unknown
331 <220> FEATURE:
332 <221> NAME/KEY: modified_base
333 <222> LOCATION: (9)
334 <223> OTHER INFORMATION: a, c, t, g, other or unknown
336 <400> SEQUENCE: 11

W--> 337 tangcnacng gcatgatatt 20
340 <210> SEQ ID NO: 12

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/004,115A

DATE: 11/06/2002
TIME: 14:53:56

Input Set : A:\73727249.app
Output Set: N:\CRF4\11062002\J004115A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; Xaa Pos. 6,7
Seq#:9; N Pos. 3,6,9
Seq#:10; N Pos. 3,6,9
Seq#:11; N Pos. 3,6,9
Seq#:12; N Pos. 3,6,9
Seq#:13; N Pos. 3,6,9
Seq#:14; N Pos. 3,6,9
Seq#:15; N Pos. 14,63,116,148,185,235,346,460,463,485,489,491,520,524,593
Seq#:15; N Pos. 602,682
Seq#:18; N Pos. 14
Seq#:19; N Pos. 434,440,443,448,458,467,475,506,510,517,522,533,535,583,584
Seq#:19; N Pos. 607,616,634,639,648,651,652,659,662,664,670,671,690,693,695
Seq#:19; N Pos. 702,717,723,728,734

VERIFICATION SUMMARY

DATE: 11/06/2002

PATENT APPLICATION: US/10/004,115A

TIME: 14:53:56

Input Set : A:\73727249.app

Output Set: N:\CRF4\11062002\J004115A.raw

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:223 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (5) SEQUENCE:
L:237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:244 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (7) SEQUENCE:
L:283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:310 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
L:337 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:391 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:418 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
L:512 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:513 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:60
L:514 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:120
L:515 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:180
L:517 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:300
L:519 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:420
L:520 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:480
L:521 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:540
L:522 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:600
L:523 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:660
L:561 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0
L:742 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:420
L:743 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:480
L:744 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:540
L:745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:600
L:746 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:660
L:747 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:720
L:1038 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (31) SEQUENCE: